

双目立体视觉的三维人脸重建方法

贾贝贝, 阮秋琦

(北京交通大学 计算机与信息技术学院, 北京 100044)

摘要:创建逼真的三维人脸模型始终是一个极具挑战性的课题. 随着三维人脸模型在虚拟现实、视频监控、三维动画、人脸识别等领域的广泛应用, 三维人脸重建成为计算机图像学和计算机视觉领域的一个研究热点. 针对这一问题, 提出一种基于双目立体视觉的三维人脸重建方法, 重建过程中无需三维激光扫描仪和通用人脸模型. 首先利用标定的2台摄像机获取人脸正面图像对, 通过图像校正使图像对的极线对齐并且补偿摄像机镜头的畸变; 在立体匹配方面, 选择具有准确可靠视差的人脸边缘特征点作为种子像素, 以种子像素的视差作为区域生长的视差, 在外极线约束、单调性约束以及对应匹配的边缘特征点的约束下, 进行水平扫描线上的区域生长, 从而得到整个人脸区域的视差图, 提高了对应点匹配的速度和准确度; 最后, 根据摄像机标定结果和立体匹配生成的视差图计算人脸空间散乱点的三维坐标, 对人脸的三维点云进行三角剖分、网格细分和光顺处理. 实验结果表明, 该方法能够生成光滑、逼真的三维人脸模型, 证明了该算法的有效性.

关键词:三维人脸模型; 双目立体视觉; 摄像机标定; 极线几何; 立体匹配; 区域生长; 三角剖分

中图分类号:TP391 **文献标识码:**A **文章编号:**1673-4785(2009)06-0513-08

3D face reconstruction using binocular stereo vision

JIA Bei-bei, RUAN Qiu-qi

(School of Computer Science and Information Technology, Beijing Jiaotong University, Beijing 100044, China)

Abstract: Generation of lifelike 3D human faces is a challenging task. Recent increases in the use of 3D face models in virtual reality, video surveillance, 3D animation, and face recognition have led to 3D face reconstruction becoming a research hotspot. The authors proposed a 3D face reconstruction method based on binocular stereo vision theory. After capturing the front view of a face with two calibrated cameras, the captured pair of stereo images were rectified to align their epipolar lines and compensate for image distortions. To obtain maps with accurate matching and dense disparity, a stereo matching algorithm based on region growing was developed. An edge feature point with reliable disparity was selected as a seed point. Then region growing was performed along horizontal scan-lines under multi-constraints. With camera calibration and disparity mapping, 3D coordinates of corresponding points were calculated. Then the face model was reconstructed with a series of methods, such as Delaunay triangulation, mesh subdivision and smoothing. Experimental results showed that the method can generate a smooth and lifelike 3D face model.

Keywords: 3D face model; binocular stereo vision; camera calibration; epipolar geometry; stereo matching; region growing; delaunay triangulation

由于人脸信息在人类日常生活和生产实践中起着十分重要的作用, 而人脸重建提供了表达和处理人脸信息的有效手段, 因此三维人脸重建是一项具

有重大意义的研究课题. 自20世纪70年代 Parke 建立第一个脸部模型^[1]开始, 许多研究人员致力于三维人脸建模的研究. 近年来, 随着三维人脸模型在虚拟现实、三维动画、视频监控、人脸识别等领域的广泛应用, 三维人脸重建已经成为计算机图像和计算机视觉领域的一个研究热点. 然而, 由于人脸具有复杂的几何形状和表面材质, 同时头发的模拟、逼真

收稿日期: 2009-04-15.

基金项目: 国家自然科学基金资助项目(60672062); 高等学校博士学科点专项科研基金资助项目(200800040008).

通信作者: 贾贝贝. E-mail: jiabeibei0129@163.com.

纹理信息的添加、光照的处理等诸多问题都使得三维人脸建模始终是一个极具挑战性的课题。

本文以双目立体视觉理论为基础,以人脸为研究对象,构建一个基于双目立体视觉的三维人脸重建系统,重建过程中无需三维激光扫描仪和通用人脸模型。首先进行摄像机标定,利用双目摄像机采集左右人脸图像对,校正图像对使极线对齐,然后选择具有准确可靠视差的人脸边缘特征点作为种子像素,以种子像素的视差作为区域生长的视差,采用基于区域生长的立体匹配算法获取整个人脸的视差图进而计算人脸三维点云数据,最后对三维点云进行三角剖分、网格细分和光顺处理重建光滑逼真的三维人脸模型。

1 三维人脸重建研究现状

获取脸部的三维数据是形成特定人脸模型的基础,根据三维数据的来源不同,国内外现有的三维人脸重建方法可分为2类:基于三维数据的三维人脸重建方法和基于图像的三维人脸重建方法。

第1类是基于三维数据的三维人脸建模方法,即从激光扫描仪(laser scanner)^[2-3]得到三维数据进行模型的拟合。Lee等人^[2]利用通过激光扫描仪获得的三维几何数据和纹理信息,调整已有的标准网格头部模型来构建特定人脸模型;Blanz和Vetter^[3]利用激光扫描仪建立了一个包括三维距离信息和表面纹理信息的头部数据库。利用激光扫描获取的三维人脸数据精度高、速度快,但是由于设备是侵入式的且价格昂贵,所以实用性较差。

第2类是基于图像的三维人脸建模方法,即从普通摄像机获取的一幅或多幅图像出发进行建模,比较有代表性的有修改标准模型和采用双目立体视觉的方法。修改标准模型进行三维人脸建模是指将标准人脸模型作为先验知识,通过刚体变换、形变等手段与具体的人脸数据拟合达到形状匹配,并且利用图像合成面部纹理信息,从而建立逼真的三维人脸模型。Ansari等人^[4]从正面和侧面人脸图像提取信息进行标准人脸模型的修正,但需要严格侧面人脸图像,否则算法失败的可能性较大。Liu等人^[5]提出利用MBA(model-based bundle adjustment)方法形成特定人脸模型,但需要人工标注特征点。

通过分析可以看出,上述的三维人脸重建方法都存在各自的局限性。而基于双目立体视觉的三维人脸重建方法^[6-7]从人类视觉机理角度出发,无需

昂贵的三维扫描仪和标准人脸模型,采用2台摄像机同时从不同角度拍摄人脸,进行立体图像匹配,再采用三角测量的方法从视差中恢复三维人脸数据。由于该方法采用的设备简单,而且可以实现实时采集,现在已成为机器视觉研究领域的一个研究热点。

2 摄像机标定

要从立体人脸图像对中提取人脸的三维信息,必须要进行摄像机的标定^[8],即根据一组已知的条件利用摄像机模型来获取摄像机的各个内外参数。比较常用的摄像机模型是针孔透视模型。于是,得到像平面坐标系与世界坐标系的映射关系:

$$Z_c = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}. \quad (1)$$

式中:\$(u, v)\$和\$(X_w, Y_w, Z_w)\$为空间\$P\$点在图像坐标系和世界坐标系下的坐标,\$f_x, f_y, u_0, v_0\$是摄像机内部参数,旋转矩阵\$R\$与平移向量\$t\$描述摄像机与世界坐标系之间的关系,称之为摄像机的外部参数。

这里,采用基于平面棋盘格标记物的摄像机标定方法。为简化计算,设模板所在平面为世界坐标系的\$z_w=0\$平面。用\$\mathbf{r}_i\$表示\$R\$的第\$i\$列向量,那么对于模板平面上的点都有:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = A [\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{r}_3 \quad t] \begin{bmatrix} X_w \\ Y_w \\ 0 \\ 1 \end{bmatrix} = A [\mathbf{r}_1 \quad \mathbf{r}_2 \quad t] \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix}. \quad (2)$$

令

$$\begin{aligned} \bar{\mathbf{M}} &= [X_w \quad Y_w \quad 1]^T, \\ \bar{\mathbf{m}} &= [u \quad v \quad 1]^T, \\ \mathbf{H} &= A [\mathbf{r}_1 \quad \mathbf{r}_2 \quad t] = [\mathbf{h}_1 \quad \mathbf{h}_2 \quad \mathbf{h}_3]. \end{aligned}$$

则式(2)可写为

$$s\bar{\mathbf{m}} = \mathbf{H}\bar{\mathbf{M}}. \quad (3)$$

式中:\$A\$是摄像机内部参数矩阵,是模板上的点和它的像点之间的一个映射。

若已知模板点的空间坐标和图像坐标,通过解最小二乘方程并用Levenberg-Marquarat算法进一步求精,就可以求得\$H\$矩阵,据此求摄像机的内外参数。利用Choleski分解就可以先得到内参,然后利用

矩阵 A 和 H , 求摄像头外参数, 如式(4)。

$$\begin{cases} r_1 = \lambda A^{-1} h_1, r_2 = \lambda A^{-1} h_2, r_3 = r_1 \times r_2, \\ t = \lambda A^{-1} h_3, \lambda = \frac{1}{\|A^{-1} h_1\|} = \frac{1}{\|A^{-1} h_2\|}. \end{cases} \quad (4)$$

图像畸变会对标定结果产生较大的影响。一般地, 标定主要考虑径向畸变, 忽略其他畸变因素。这里, 采用 Zhang^[9] 的方法来处理畸变问题。

3 立体匹配算法

基于双目立体视觉的三维人脸重建系统的一个关键技术是求解立体图像对的对应点问题, 即解决立体匹配问题, 以便得到视差图 (disparity map)。目前, 立体匹配算法^[10] 主要分为基于特征和基于区域2类。前一类选取图像中具有惟一性质的点作为匹配点, 通常选择边缘、角点等, 大大减少了搜索范围, 但是只能得到稀疏的视差图; 后一类考虑每个需匹配点的邻域性质, 能够得到稠密的视差图, 但是对无纹理的区域表现欠佳, 且计算量大, 匹配速度慢。通过以上分析, 采用结合特征技术的基于区域生长算法的立体匹配算法。

3.1 基于区域生长的立体匹配算法

区域生长算法^[11] 的基本思想是将具有相似性质的像素集合起来构成区域。首先, 在每个需要分割的区域选择一个种子像素作为生长的起始点, 然后根据某种事先确定的生长或相似准则, 将种子像素周围邻域中与种子像素有相同或相似性质的像素合并到种子像素所在的区域。此时将新合并的像素作为新的种子像素, 继续以上搜索和合并过程, 直到没有可以合并的像素为止, 这样一个区域就生长成了。

如图1所示, 若已知左图像上的一个种子像素 L_1 与右图像上一点 R_1 匹配, 则与 L_1 相邻的点 L_2 和 L_3 的匹配点必定在 R_1 点的附近。这样 L_2 和 L_3 的匹配点的搜索范围就可以限制在 R_1 点的一个较小的邻域内。通常只需要在十几个像素大小的搜索窗内计算匹配代价函数, 求出最大的匹配度, 即可认为是所求点的可靠匹配。这里, 采用差值平方和 (SSD) 定义匹配代价函数 $e(i, j, d)$:

$$e(i, j, d) = \frac{1}{mn} \sum_{x=i}^{i+m} \sum_{y=j}^{j+n} (I_L(x, y) - I_R(x, y + d))^2. \quad (5)$$

式中: $I_L(x, y)$ 和 $I_R(x, y)$ 分别为左图像和右图像第 x 行、第 y 列像素的灰度值, d 为视差, $m \times n$ 为模板尺寸。选择左图像为参考图像, 右图像为匹配图像,

则位于左图像 (x, y) 处的像素和位于右图像 $(x, y + d)$ 处的像素是对应点。

与传统的逐点计算的方法相比, 区域生长算法极大提高了匹配的效率和匹配效果也更好。

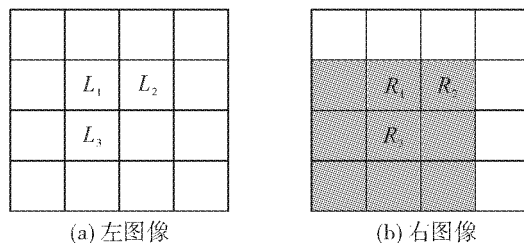


图1 种子像素确定的区域生长范围

Fig. 1 The range of region growing by seed points

3.2 改进的基于区域生长的立体匹配算法

3.2.1 约束条件

1) 极线约束: 如图2所示, O_L 和 O_R 分别为左、右像平面的光心, p 和 p' 分别为空间一点 P 点在左、右像平面的投影点, 平面 $PO_L O_R$ 成为极平面, 极平面与左、右像平面的交线 pe 和 $p'e'$ 成为极线。极线约束了立体图像对应点的位置, 即空间点 P 在左像平面上投影点 p 所对应的右像平面投影点必在极线 $p'e'$ 上, 反之亦然。

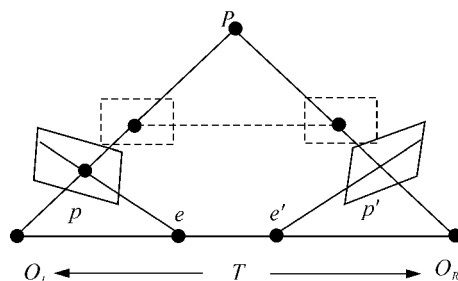


图2 极平面和极线示意图

Fig. 2 Epipolar plane and epipolar lines

2) 单调性约束: 沿着匹配点的共轭极线, 所有可见点都保持同样的次序, 即匹配点在2幅图像中的先后顺序必须一致。根据单调性约束, 区域内每个像素的匹配点就限制在右图像对应像素行的一定区域内。

3) 平行立体视觉: 为了进一步降低算法的复杂度, 双目系统采用的是平行光轴模型, 即平行双目立体视觉系统, 此时外极线与图像扫描线是重合的, 如图2中虚线所示。所以块匹配的搜索窗由二维区域变为一维线段, 匹配点只需要在扫描线上寻找。依次对每条水平扫描线上的像点在极线约束和单调性约束下进行匹配, 进而得到立体图对的视差。因此, 只需要在一维水平方向上进行区域生长, 即只需要考虑与种

子像素相邻的像素点,这大大缩小了搜索空间。

3.2.2 种子像素的初始化

基于区域生长的立体匹配算法^[12]的关键是需要找到少量准确、可靠的匹配点对作为种子像素,之后就可以根据它们之间的相对位置关系将匹配迅速传播到整个图像区域,这对于提高系统的可靠性和效率有着直接的影响。考虑到三维重建的对象是人脸,首先要确定人脸的区域,而图像中人脸的边缘特征是三维人脸重建的重要信息,它在2幅图像中产生相似的结果并且分布较合理,所以选择人脸边缘特征点作为初始种子像素。

选择用 Canny 算子提取单像素人脸边缘,根据以上约束条件在相应的水平扫描线上进行边缘特征点的匹配。对左图像的每个特征点,把其在右图像相应的像素行中所有可能的匹配点组成一个集合。其中 d 或者是特征点与可能匹配点的视差,或者是代表无匹配点的特殊符号。对每个可能匹配点,可以设定初始匹配概率 $P^{(0)}(l)$, 计算式(5)的匹配代价函数, $e(i, j, d)$ 与初始匹配概率 $P^{(0)}(l)$ 成反比。换句话说, $P^{(0)}(l)$ 与可能匹配点邻域中的相似度有关。因此,可以利用松弛迭代法,给可能匹配点邻域中视差比较接近的点以正的增量,而给可能匹配点邻域中视差比较远的点以负的增量来对 $P^{(0)}(l)$ 进行迭代更新。随着迭代的进行,正确匹配点的第 k 次迭代匹配概率 $P^{(k)}(l)$ 会逐渐增大,而其他点的匹配概率 $P^{(k)}(l)$ 会逐渐减小,经过一定次数的迭代后,将匹配概率 $P^{(k)}(l)$ 最大的点,也就是匹配代价最小的点确定为匹配点,此时的 d 就是左图像点 (i, j) 处匹配可信度最高的视差估计。

3.2.3 改进的区域生长算法

基于区域生长的立体匹配算法的2个关键是:

- 1) 选择一组能正确代表所选区域的种子像素;
- 2) 确定在生长过程中能将相邻像素包括进来的准则。

这里设定这样的准则:对于与种子像素相邻的点,将种子像素的视差值 d 代入匹配代价函数,如果满足 $e(i, j, d) < T$, 其中 T 为预先给定的匹配代价门限,就将相邻像素包括进来,这时合并到种子像素所在区域的所有像素具有相同的视差值,即整个生长区域具有相同的视差值。上述准则可以描述为

$$\begin{cases} d(i+1, j) = d(i, j), e(i+1, j, d(i, j)) \leq T; \\ d(i+1, j) = d_{\min\{e(i+1, j, d)\}}, e(i+1, j, d(i, j)) > T. \end{cases} \quad (6)$$

实际上,该算法可以理解为基于视差生长的立体匹配算法,具体算法如下:

1) 种子像素选取:初始化时,选择边缘特征点作为种子像素进行区域生长。当以边缘特征点作为种子像素的区域生长完成后,选择一个不属于任何已生长区域的像素点,在边缘特征点的视差约束下利用式(5)计算一维的搜索窗中的匹配代价函数,选取最可靠的视差。选取该点为种子像素,并且设该点视差为区域视差,然后执行2)。如果没有找到任何小于预先给定的匹配代价门限 T 的匹配点,对下一个相邻像素点重复这个步骤。

2) 区域生长:利用种子像素的视差值,计算与种子像素相邻的像素点的匹配代价。如果满足式(6),则将该像素包括进种子像素所在区域,否则丢弃该点。

3) 视差图生成:反复执行2)直到没有可以再合并的像素为止,这样一个区域就生长成了。返回1)找到新的种子像素后重复以上步骤。当图像中所有像素点被标记过后,则结束算法。此时,视差图 $d(i, j)$ 生成。

4 三维人脸重建算法

4.1 空间点的三维重建算法

本文采用的双目立体视觉系统由2个相同的摄像机组成,且它们坐标系统的各对应轴平行,即平行光轴模型。设摄像机的焦距为 f , 2个光心间的距离(基线)为 b , 以左摄像机坐标系为世界坐标系,空间点 P 的世界坐标为 (X, Y, Z) , 在左右摄像机坐标系下的坐标分别为 (x_1, y_1, z_1) 和 (x_2, y_2, z_2) , 在左右图像平面上的坐标为 (u_1, v_1) 和 (u_2, v_2) , 其中视差 $d = u_1 - u_2$ 。

根据成像的几何比例关系,对于左右摄像机有如下关系:

$$\begin{cases} \frac{f}{z_1} = \frac{u_1}{x_1} = \frac{v_1}{y_1}, \\ \frac{f}{z_2} = \frac{u_2}{x_2} = \frac{v_2}{y_2}. \end{cases} \quad (7)$$

世界坐标系与摄像机坐标系的关系表述如下:

$$\begin{cases} X = x_1 = x_2 + b, \\ Y = y_1 = y_2, \\ Z = z_1 = z_2. \end{cases} \quad (8)$$

由此,可计算空间点的三维坐标:

$$\begin{cases} X = \frac{Z}{f}u_1 = \frac{u_1 \cdot b}{d}, \\ Y = \frac{Z}{f}v_1 = \frac{v_1 \cdot b}{d}, \\ Z = \frac{f \cdot b}{d}. \end{cases} \quad (9)$$

上式把三维物体的深度 Z 与视差 d 直接联系起来,而深度反映了三维空间信息.如果已知基线和焦距,确定视差 d 后就可以计算 P 点的 Z 坐标,进而求出 P 点的世界坐标 X 和 Y .这样,运用式(9)就可以得到人脸立体图像对所有匹配点的三维坐标.为了得到更加光滑的三维人脸深度,在计算深度 Z 之前使用 5×5 的中值滤波器对视差 d 进行平滑处理.

4.2 三维人脸模型的三角网格生成算法

当所有匹配点的三维坐标都计算出来后,可以将这些离散点在三维空间中直接显示出来,但是为了得到真实感的三维人脸模型,有必要进行三维数据点的三角剖分.如果直接在三维空间中进行三角剖分,那么必须要讨论空间3个相邻点的位置问题,即以这3个点所构成的球体中不能包括其他的点.考虑到这种情况比较复杂,先将三维数据点一对一地投影到二维平面区域,在二维区域上进行投影点的三角化,然后通过点的对应关系基本上实现空间点的三角剖分.对二维空间离散点集的剖分而言,存在着许多种三角剖分算法.俄国数学家 Delaunay 在 1934 年证明了必定存在且仅存在一种剖分算法,使得所有三角形内角之和最大(称为 Delaunay 三角剖分).Delaunay 三角剖分具有平均形态比最大的性质,能够尽可能地避免病态三角形的出现,是目前公认的一种最佳的三角剖分方法^[12-13].

为了得到高质量三维人脸模型的三角形网格,对人脸散乱数据点进行如下操作:

1) 初始网格生成.对散乱点进行排序,搜索 X 坐标最小的点,设该点为 v_1 ,按照与 v_1 点的距离的平方递增的顺序排列各点,形成序列 v_1, v_2, \dots, v_n .将 v_1 与 v_2 相连构造第一条边,在 v_i 序列中顺序搜索与 v_1 和 v_2 不共线的点,记作 v_k ,则将 v_k 插入 v_3 前,其余点顺序后移,将 v_1, v_2, v_k 3 点相连形成第 1 个三角形和初始网格前沿边界.接下来采用网格前沿技术,逐点向外扩展,根据最小内角最大准则形成初始的人脸三角网格.

2) 网格细分.采用 Loop 细分方法^[14],这是一种

基于三角形控制网格的细分,该细分方法所生成的曲面是基于四次 B 样条曲面.现已证明细分曲面在正则点处达到 C^2 连续,在奇异点处达到 C^1 连续,具有细分规则简单、细分后光滑性好等特点. Loop 细分算法是逼近型 1-4 分裂算法,基本思想是在三角形的每一边上新插入一个结点后将其分成 2 段,原三角形的顶点将被新的结点代替,于是一个三角形单元将被 4 个小三角形单元代替.插入点生成规则如下:

①若内部边有 2 个顶点 v_0 和 v_1 ,共享此边的 2 个三角形为 (v_0, v_1, v_2) 和 (v_0, v_1, v_3) ,则新边点 v_E 为

$$v_E = \frac{3}{8}(v_0 + v_1) + \frac{1}{8}(v_2 + v_3). \quad (10)$$

②若内部顶点 v 的 1-邻域顶点为 $v_i (i = 0, 1, \dots, n-1)$,则新生成的顶点 v_V 为

$$v_V = (1 - n\beta)v + \beta \sum_{i=0}^{n-1} v_i. \quad (11)$$

式中: β 为邻点权值.

$$\beta = \frac{1}{n} \left(\frac{5}{8} - \left(\frac{3}{8} + \frac{1}{4} \cos \frac{2\pi}{n} \right)^2 \right); n = |v|_E.$$

③若边界边的 2 个顶点为 v_0 和 v_1 ,则新边点 v_E 为

$$v_E = \frac{1}{2}(v_0 + v_1). \quad (12)$$

④若边界顶点 v 在边界上的 2 相邻顶点为 v_0 和 v_1 ,则新生成的顶点 v_V 为

$$v_V = \frac{1}{8}v_0 + \frac{1}{8}v_1 + \frac{3}{4}v. \quad (13)$$

根据插入点的生成规则计算出网格每一条边的边点 v_E 和每一个顶点的顶点 v_V ,将新边点和新顶点连接起来,生成一个新的三角网格,直到最后收敛于极限曲面.

3) 网格优化:采用 Laplacian^[15] 光顺技术,通过调整三角形网格的顶点位置来提高网格质量,得到优化的三维人脸模型.设三角形网格顶点 $v_i, v_{i,j} (j = 0, 1, \dots, n-1)$ 是与 v_i 邻接的 n 个网格顶点.顶点 v_i 的一阶、二阶 Laplacian 算子定义为

$$\begin{aligned} L(v_i) &= \frac{1}{n} \sum_{j=0}^{n-1} v_{i,j}, \\ L^2(v_i) &= \frac{1}{n} \sum_{j=0}^{n-1} (v_{i,j} - L(v_{i,j})). \end{aligned} \quad (14)$$

则顶点 v_i 的新位置为

$$v'_i = v_i - \frac{1}{\gamma + 1} L^2(v_i). \quad (15)$$

式中: $\gamma = \frac{1}{n_i} \sum_j \frac{1}{n_{i,j}}$, 这里 n_i 和 $n_{i,j}$ 分别是顶点 v_i 和它的第 j 个相邻点 $v_{i,j}$ 的度.

5 实验结果与分析

5.1 基于双目立体视觉的三维人脸重建系统构成

本文的研究是以双目立体视觉理论为基础,以人脸为研究对象,构建一个三维人脸重建系统. 该系统的实验平台如图3所示,主要是由2台 JVC 彩色视频摄像机 TK-C1481BEC、大恒图像采集卡 DH-QP300、视频电缆、黑白棋盘格标定板和 Pentium 4 CPU 2.93GHz 计算机组成.

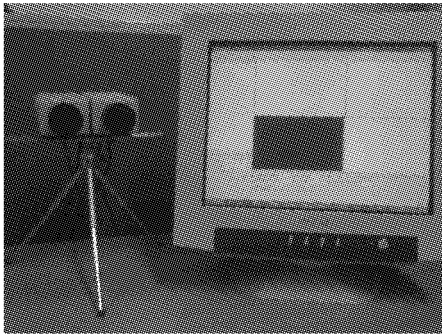


图3 三维人脸重建系统实验平台

Fig. 3 Experimental platform of 3D face reconstruction system

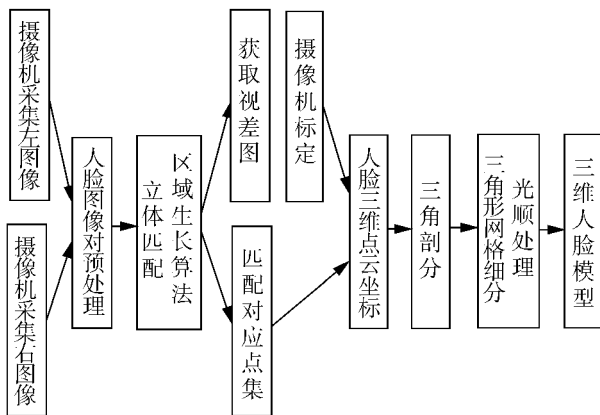


图4 三维人脸重建系统流程图

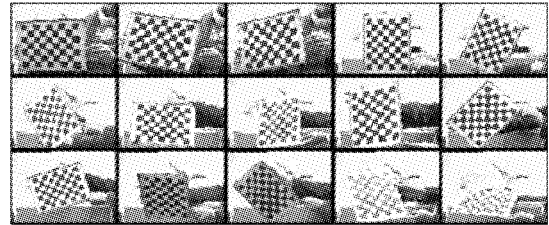
Fig. 4 Flow chart of 3D face reconstruction system

图4所示为整个系统的流程图. 首先利用标定的2台摄像机获取人脸正面图像对,通过图像校正使图像对的极线对齐. 在立体匹配方面,结合特征技术,采用基于区域生长的立体匹配算法得到整个人脸区域的视差图. 最后,将人脸的三维点云在二维平面上进行 Delaunay 三角剖分、网格细分和光顺处理来重建三维人脸模型.

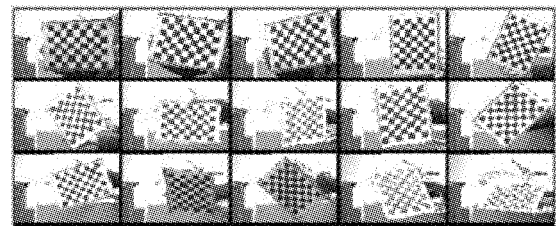
5.2 摄像机标定分析

利用上面介绍的标定方法对双目摄像机进行内外参数标定,具体实现如下:

1)左右摄像机采集标定模板图像:选择 7×9 的平面棋盘格作为标定物,棋盘格每个小格的尺寸为 $2.8 \text{ cm} \times 2.8 \text{ cm}$. 图5为用左右摄像机同时从不同方位采集到的15幅标定模板图像.



(a) 左摄像机采集15副标定模板图像



(b) 右摄像机采集15副标定模板图像

图5 双目摄像机采集标定模板图像

Fig. 5 Plane model images captured by two cameras

2)角点检测:对棋盘格进行角点检测,目的是为了得到角点的二维图像坐标,作为标定过程的数据输入,检测结果精确程度直接影响了整个系统的标定精度. 所以,采用的是精度很高的基于 Harris 算子的亚像素级角点检测算法. 图6为角点检测的结果.

3)内外参数标定:由于所采用的双目摄像机是固定焦距的,所以在使用过程中其内部参数不会发生变化. 由上述摄像机标定算法得到内外参数标定结果见表1. 可见,文中采用的立体视觉系统可以认为是由2个完全相同的摄像机组成的平行双目立体视觉系统.

表1 摄像机标定结果

Table 1 Camera calibration results

摄像机内外参数		左摄像机	右摄像机
内参数	焦距	2 092.486 47	2 082.291 71
	主点	99.229 78	169.220 02
	畸变	$\begin{bmatrix} -0.143 27 \\ -1.336 91 \\ 0.006 49 \\ 0.004 04 \\ 0.000 00 \end{bmatrix}$	$\begin{bmatrix} 0.331 21 \\ -1.055 71 \\ 0.002 95 \\ 0.001 10 \\ 0.000 00 \end{bmatrix}$
外参数	旋转向量	$[0.039 56 \ 0.023 86 \ -0.011 95]$	
	平移向量	$[75.846 07 \ 1.170 52 \ -2.954 48]$	

4)误差分析:为了表征摄像机标定的测量精度并对误差进行定量分析,引入了平均绝对误差 E_w , 定义如下:

$$E_w = \frac{\sum_{i=1}^N \sqrt{(X_{ti} - X_i)^2 + (Y_{ti} - Y_i)^2}}{N}. \quad (16)$$

式中: (X_i, Y_i) 是棋盘格角点的实际图像坐标, (X_{ti}, Y_{ti}) 是通过标定的双目视觉测量出来的图像坐标, N 为测量角点的数目. 将角点坐标重新投影到图像中, 所得误差为 $E_w = 0.15927$.

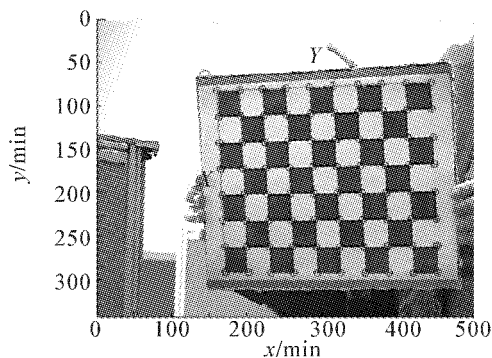


图6 角点检测

Fig. 6 Corner detection

图7为经过角点畸变校正后的重投影误差分布图. 从图中可以看出, 经角点校正后的图像误差分布均匀, 在 x 和 y 方向上检测精度均不超过0.4个像素. 因此从实验结果来看, 本摄像机标定方法是有效的.

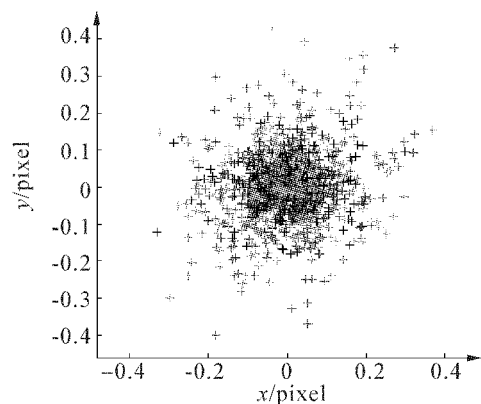


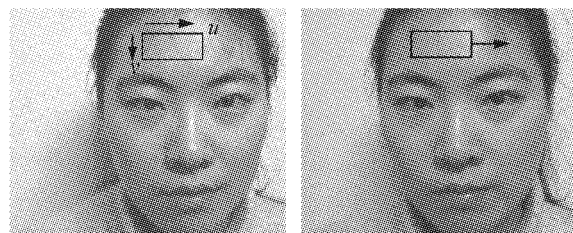
图7 重投影误差分析

Fig. 7 Reprojection error analysis

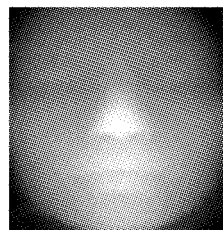
5.3 立体图像匹配和视差图生成

图8给出了立体匹配生成视差图. 图8(a)为双目摄像机拍摄的人脸图像对, 图像分辨率为 504×342 . 首先利用基于特征技术得到人脸边缘特征点3134个, 对边缘特征点做灰度相关处理, 得到准确可靠的视差. 然后选择匹配的边缘特征点2887个作为种子像素, 以种子像素的视差作为区域生长的视差, 在外极线约束、单调性约束以及对应匹配的边缘特征点的约束下, 对与种子像素相邻的像素点在水平扫描线上进行区域生长, 最终选取20329个种子像素从而得到整个人脸区域的视差图. 将生成的人脸视差图量化到256(0~255)个灰度级, 如图8(b)所示. 从图中可

以看出, 该立体匹配算法获取的人脸视差图比较光滑且能够反映出人脸较为明显的视差阶跃.



(a) 采集的人脸图像对



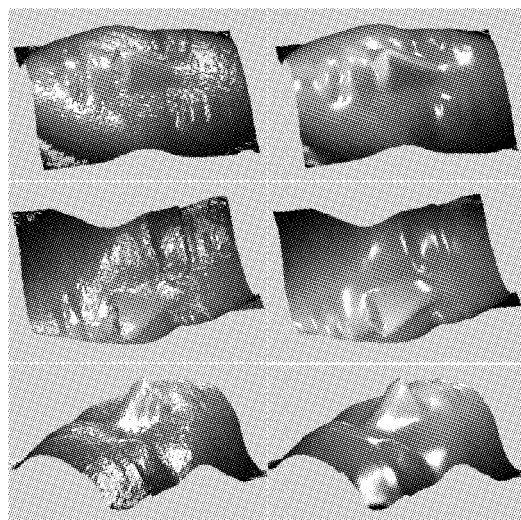
(b) 区域生长算法生成视差图

图8 立体匹配生成视差图

Fig. 8 Stereo matching and disparity map

5.4 三维人脸重建结果

采用上面介绍的三维人脸重建方法, 即利用人脸视差图和摄像机内外参数的结果计算人脸点云的三维坐标. 采用 Delaunay 方法对散乱的三维点云进行三角剖分, 如图9(a). 而后对初始生成的三角形网格三维人脸模型进行了网格细化和光顺处理, 得到最终的三维人脸模型如图9(b)所示. 从图中可以看出, 初始生成的三维人脸模型粗糙而不平滑, 经过网格优化后, 对初始人脸模型进行了孔洞修补和平滑处理, 使得三维人脸主体部分的重建效果真实感强.



(a) 初始生成三角形网格 (b) 最终的三维人脸模型的三维人脸模型

图9 多视角下的三维人脸重建结果

Fig. 9 3D face reconstruction results from multiple views

6 结束语

本文从人类视觉机理角度出发,提出一种基于双目立体视觉的三维人脸重建方法.首先利用标定的2台摄像机获取人脸正面图像对,通过图像校正使图像对的极线对齐,然后结合特征技术,选择具有准确可靠视差的人脸边缘特征点作为种子像素,以种子像素的视差作为区域生长的视差,在水平扫描线上进行区域生长,提高了匹配的速度和准确度,从而得到整个人脸区域的视差图.最后根据摄像机标定和立体匹配结果计算得到空间点的三维坐标,通过对人脸的三维点云进行 Delaunay三角剖分、网格细分和光顺处理,得到光滑逼真的三维人脸模型.

在今后的研究中,主要解决由于光照因素对三维人脸重建造成的影响,考虑采用主动光源对被拍摄的人脸区域进行照射,同时进行人脸图像对的采集,这样可以有效地减小光线变化对人脸图像的影响,从而达到在各种光照条件下重建出具有真实感的三维人脸模型.

参考文献:

- [1] PARKE F I. Computer generated animation of faces[C]//Proceedings of ACM Annual Conference. Boston, United States, 1972: 451-457.
- [2] LEE Y C, TERZOPOULOS D, WATERS K. Realistic face modeling for animation[C]//SIGGRAPH Proceeding. Los Angeles, USA, 1995: 55-62.
- [3] BLANZ V, VETTER T. A morphable model for the synthesis of 3D faces[C]//SIGGRAPH Proceedings. Orlando, FL, USA, 1999: 187-194.
- [4] ANSARI A, ABDEL-MOTTALEB M. 3-D face modeling using two views and a generic face model with application to 3-D face recognition[C]//Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance. Miami, USA, 2003: 37-44.
- [5] LU Z, ZHANG Z, JACOBS C, et al. Rapid modeling of animated faces from video M SR-TR-2000-11[R]. 2000.
- [6] 王宣银, 潘峰, 向桂山, 等. 基于Snake模型的特定人脸三维重建方法[J]. 机械工业学报, 2007, 43(7): 168-173.
- WANG Xuanyin, PAN Feng, XIANG Guishan, et al. Specific face 3D reconstruction based on Snake model[J]. Chinese Journal of Mechanical Engineering, 2007, 43(7): 168-173.
- [7] CHEN Q, MEDIONI G. Building 3-D human face models from two photographs[J]. The Journal of VLSI Signal Processing, 2001, 27(2): 127-140.
- [8] 张广军. 机器视觉[M]. 北京: 科学出版社, 2005: 58-84.
- [9] ZHANG Zhengyou. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.
- [10] MYRON Z, DARIUS BURSCHKA, GREGORY D. Advances in computational stereo[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(8): 993-1008.
- [11] 阮秋琦. 数字图像处理学[M]. 2版. 北京: 电子工业出版社, 2007: 382-408.
- [12] SHEWCHUK JR. Delaunay refinement algorithms for triangular mesh generation[J]. Computational Geometry-Theory And applications, 2002, 22(113): 21-74.
- [13] 张永春, 达飞鹏, 宋文忠. 三维散乱点集的曲面三角剖分[J]. 中国图象图形学报, 2003, 8(12): 1379-1388.
- ZHANG Yongchun, DA Feipeng, SONG Wenzhong. Surface triangulations based on arbitrary point-sets[J]. Journal of Image and Graphics, 2003, 8(12): 1379-1388.
- [14] KOBBELT L. A variational approach subdivision[J]. Computer Aided Geometric Design, 1996, 13(8): 743-761.
- [15] FIELD D A. Laplacian smoothing and Delaunay triangulations[J]. Communications in Applied Numerical Methods, 1988, (4): 709-712.

作者简介:



贾贝贝,女,1985年生,硕士研究生,主要研究方向为计算机视觉、数字图像处理、信号与信息处理。



阮秋琦,男,1944年生,教授,博士生导师,北京交通大学信息科学研究所所长,国务院学位委员会学科评议组成员,IEEE高级会员.主要研究方向为图像处理、计算机视觉、多媒体信息处理、虚拟现实、计算机支持的协同工作.曾获国家教委科技进步二等奖、铁道部科技进步二等奖和三等奖等,获茅以升科技奖、詹天佑科技人才奖并享受国家政府津贴的荣誉.发表学术论文150余篇,出版著作3部,获国家专利1项。